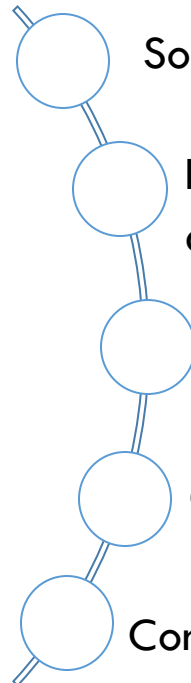


Integrazione tra indagini campionarie e archivi amministrativi per il disegno di politiche locali

Antonio Schizzerotto

Convegno Nazionale USCI
Trento, 14-15 settembre 2017

IMPORTANZA DEI DATI AMMINISTRATIVI PER IL DISEGNO DI POLITICHE SOCIALI

- 
- 1 Sono riferiti a tutta la popolazione pertinente e, di massima, correntemente aggiornati
 - 2 Il loro utilizzo comporta grandi vantaggi operativi: riduzione dei costi di rilevazione, eliminazione del disturbo statistico per i rispondenti e relativi problemi di memoria o di reticenza
 - 3 Consentono di produrre stime riferite a piccole aree
 - 4 Consentono di monitorare in maniera adeguata le politiche
 - 5 Consentono di generare basi di dati longitudinali

MA...

- ✓ richiedono completezza, affidabilità e integrabilità ai fini di una loro utilizzazione statistica
- ✓ sono carenti di informazioni che possono risultare cruciali sia per il disegno sia per la valutazione di impatto delle politiche sociali
- ✓ anche in presenza di identificativi anonimizzati si pone un problema quando gli archivi (soprattutto quelli che contengono informazioni individuali ricche) sono riferiti a individui identificabili

IMPORTANZA DEI DATI DI *SURVEY* PER IL DISEGNO DI POLITICHE SOCIALI



Ricchezza informativa (soprattutto se realizzate ad hoc per indagare uno specifico tema)

Possono raccogliere dati sull'intera famiglia dell'intervistato

Possono includere informazioni retrospettive

Possono avere un carattere longitudinale

MA...

- ✓ possono avere una contenuta dimensione campionaria
- ✓ non rilevano indicatori 'di tipo amministrativo' quali Icef, necessari per identificare gli ammissibili ad alcuni interventi
- ✓ affidabilità di informazioni su redditi e patrimoni (errate dichiarazioni, effetto memoria)
- ✓ reticenza degli intervistati a rispondere ad alcune domande
- ✓ esigenze di semplificazione in presenza di obiettivi conoscitivi ampi

UNA POSSIBILE STRATEGIA: L'INTEGRAZIONE DI BASI DI DATI DA UNA PLURALITÀ DI FONTI

L'esperienza di utilizzo integrato di dati da varie fonti – di origine amministrativa e da surveys – può essere condotta a due livelli:

confrontando stime di aggregati e/o indicatori da fonti diverse, soprattutto a fini di raccordo di serie (per ovviare alla scarsa tempestività di una delle due fonti);

procedendo all'**abbinamento** e all'**integrazione** di molteplici registri amministrativi e degli stessi con microdati da surveys

Tramite **matching esatto**,

con chiave il codice fiscale (necessario attuare procedure per tutela della *privacy*: eliminazione dell'identificatore diretto, aggregazione di modalità/valori di variabili con frequenze inferiori a 3, soppressione del giorno/mese di nascita)

Tramite metodi basati su **modello**,

integrazioni tra informazioni contenuti in fonti differenti rappresentanti la stessa popolazione ma non contenenti le stesse unità)

A) l'esperienza di integrazione dei microdati Inps e Indagine panel sulle condizioni di vita per l'analisi delle dinamiche del mercato del lavoro locale (via matching esatto)

Cassetto previdenziale INPS

Vantaggi:

- contiene le storie lavorative documentate (occupazione e disoccupazione sussidiata) dell'intera popolazione

Svantaggi:

- copertura parziale dell'universo (non tutti i lavoratori e non tutti gli episodi di lavoro)
- problemi di qualità o di pertinenza delle informazioni rispetto alle esigenze conoscitive
- mediocre tempestività nell'aggiornamento della fonte
- vincoli di mancata disaggregazione secondo alcuni caratteri individuali (es.territorio)
- assenza di informazioni individuali e familiari rilevanti (es.titolo di studio, origini sociali, condizioni di vita)

Indagine Panel

Vantaggi:

- contiene le storie lavorative dichiarate (occupazione, disoccupazione, inattività)
- rileva il lavoro senza contratto e il lavoro autonomo
- rileva la famiglia dell'intervistato e le condizioni di vita
- rileva la storia educativa e matrimoniale

Svantaggi:

- problemi di qualità delle informazioni rispetto alle esigenze conoscitive
- mancate risposta

Ai fini della ricostruzione delle storie lavorative, si resa necessaria una **complessa attività di affinamento dei dati amministrativi**, organizzati non per fini statistici (basti dire che, ad es. per normativa uno stesso episodio può comparire negli archivi INPS in più di una registrazione)

Le procedure di *affinamento della base di dati Inps* hanno riguardato:

- eliminazione dei record *multipli*
- unione episodi di lavoro *segmentati*
- correzione degli episodi di *mobilità spuria*

Tra le due fonti vi sono alcune importanti differenze nella mappatura della storia lavorativa:

	Panel	Inps
<i>i. Condizione di lavoro</i>	Tutti i lavori	Mancano per definizione episodi di lavoro dipendente pubblico, di libero professionista con cassa previdenziale propria e di lavoratore in nero. Parziale copertura degli agricoli e degli autonomi
<i>ii. Condizione di disoccupazione</i>	Chi si dichiara tale	Solo sussidiata
<i>iii. Cadenza temporale delle osservazioni</i>	i segmenti di storie lavorative dichiarati dagli intervistati avendo come unità di tempo il mese	episodi documentati con unità di tempo il giorno
<i>iv. Sovrapposizioni tra episodi</i>	Non ammesse perché si rileva la condizione prevalente	Ammessi

Strategie di integrazione delle due fonti:

Nell'integrare le informazioni tratte dalle due fonti si è seguito, quando possibile, un criterio di prevalenza dell'informazione amministrativa.

Per quanto riguarda le storie lavorative, la fonte *Inps* è stata trattata come fonte:

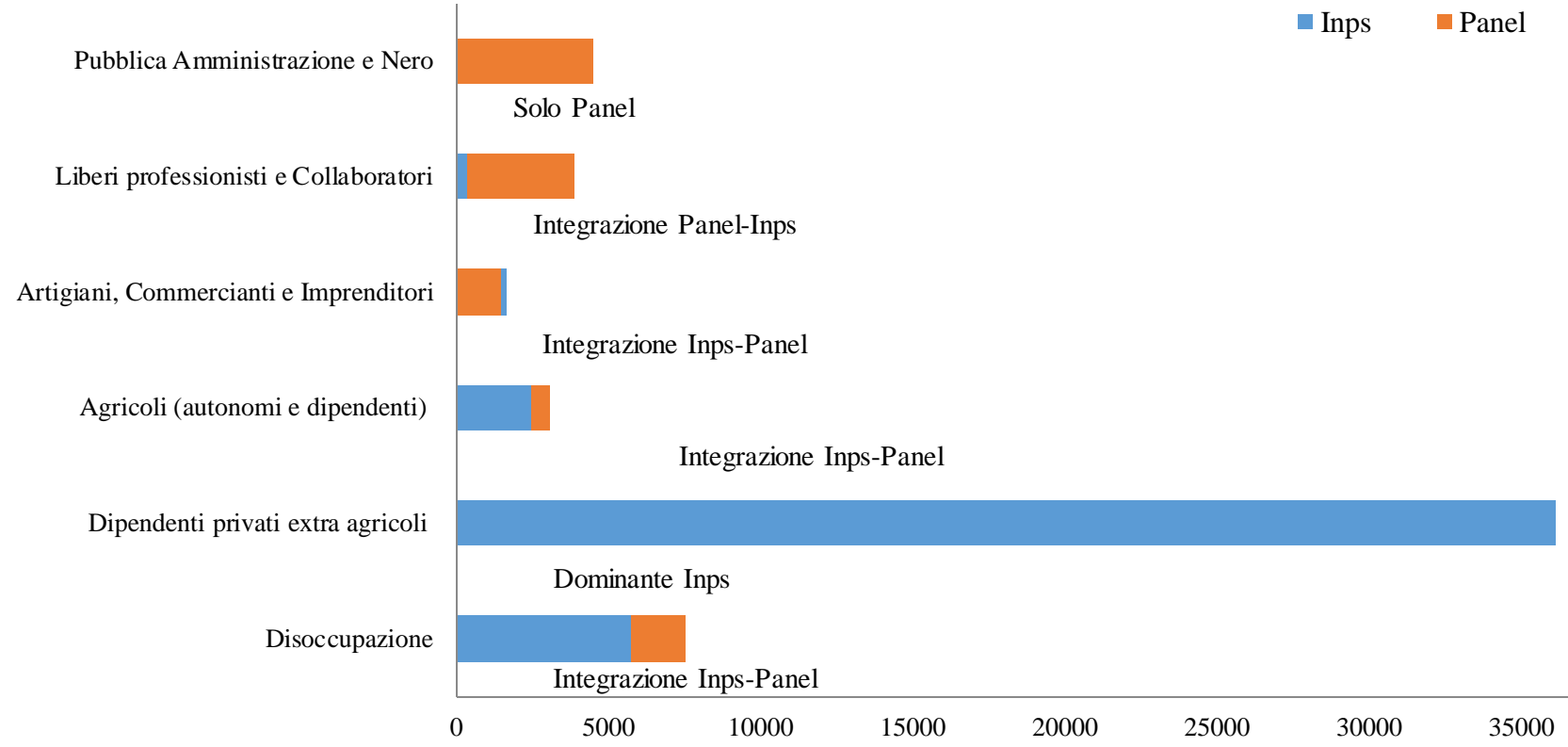
- a) **esclusiva** nel caso dei dipendenti privati extra-agricoli;
- b) **iniziale, integrata da *Panel***, nel caso di soggetti operanti nell'agricoltura (autonomi e dipendenti), artigiani e commercianti e imprenditori che non risultavano nel cassetto previdenziale Inps .

Nei restanti casi, si è fatto ricorso:

- c) al dato ***Panel* come fonte esclusiva** solo nei casi in cui un individuo avesse episodi non rinvenibili, per definizione, in esso (episodi di lavoro dipendente nelle amministrazioni pubbliche, episodi di lavoro "nero");
- d) **prioritariamente al *Panel*** per gli episodi di attività di libero professionista e di collaboratore: (d1) ai primi perché, se con cassa propria, presenti soltanto in *Panel*; (d2) ad entrambi perché rinvenibili in *Inps* con significativi limiti informativi sulle date di inizio e di fine.

Per quanto attiene, invece, gli episodi di disoccupazione, si è utilizzato il criterio di **priorità a *Inps***. Quindi, si è fatto ricorso al *Panel* solo quando non si sovrapponeva a periodi di lavoro così come sopra descritti o a periodi di disoccupazione sussidiata.

Esito delle strategie di integrazione delle due fonti: episodi di lavoro secondo la categoria occupazionale e la fonte



(B) L'esperienza di integrazione dei microdati in materia di transizioni scuola lavoro

L'integrazione, mediante matching esatto, ha coinvolto 2 archivi amministrativi e 2 dati di survey:

1) **Anagrafe** qualificati e diplomati della Formazione Professionale e maturi

Vantaggi:

- dato censuario su tutti i qualificati e i diplomati FP (dal 2006 al 2013) e i maturi (dal 2007 al 2013)

Svantaggi:

- Non è nota la storia lavorativa
- Non sono note la situazione familiare

2) **Archivio Comunicazioni Obbligatorie** (episodi di lavoro dipendente) dal 2007 al 2012 integrate con gli archivi ASIA e APIA che rilevano il lavoro autonomo e agricolo

Vantaggi:

- sequenza degli episodi lavorativi (tutta la carriera)

Svantaggi:

- non è noto lo stato di inattività dei non-occupati
- non è possibile distinguere i non-occupati dagli occupati fuori-provincia
- Assenti variabili familiari e caratteristiche individuali rilevanti (titolo di studio, origini sociali...)

3) Indagine telefonica censuaria sui qualificati e diplomati della Formazione Professionale (leva 2010) (Agenzia del Lavoro) e

4) Indagine telefonica sui maturi (leva 2006) (Agenzia del Lavoro)

Vantaggi:

- Rilevata la condizione di inattività lavorativa (al 18esimo mese dopo la fine degli studi)
- Rilevata l'informazione sul numero di mesi spesi alla ricerca attiva del primo lavoro
- Informazioni sul primo lavoro e sul lavoro al momento dell'intervista (18 mesi dopo la fine degli studi)

Svantaggi:

- tasso di caduta potenzialmente non-casuale
- impossibile distinguere qualifica e diploma professionale prima del 2010

(C) L'esperienza di integrazione dei microdati in materia di transizioni scolastiche

Anagrafe Unica degli Studenti della Provincia di Trento (AUSPAT)

Si compone di tre archivi amministrativi distinti collegabili tra loro mediante chiave identificativa univoca dello studente:

- Iscrizioni: l'archivio contiene informazioni relative ad ogni iscritto al sistema scolastico-formativo provinciale dalla scuola primaria alla scuola secondaria di secondo grado (inclusa la Formazione Professionale)
- Esami: l'archivio contiene informazioni relative agli esiti degli esami conclusivi del primo ciclo di istruzione, della formazione professionale (triennale e quadriennale) e degli esami di maturità
- Responsabili: l'archivio contiene informazioni relative ai responsabili adulti degli iscritti al sistema scolastico formativo

Vantaggi:

- L'unione degli archivi consente di indagare le determinanti della scelta dell'indirizzo formativo, in particolar modo la specifica scuola di provenienza, i voti e la regolarità scolastica e l'influenza esercitata dalle origini sociali familiari.

Svantaggi:

- disponibili per gli anni scolastici dal 2006/2007 in poi
- Non contengono gli studenti fuori provincia
- Non contengono informazioni sulla storia lavorativa

(D) L'esperienza di integrazione dei microdati per l'analisi dei corsi di formazione per disoccupati

I corsi di formazione hanno aumentato le chance occupazionali dei partecipanti? I benefici economici derivanti dalla realizzazione dei corsi hanno coperti i costi?

Si sono utilizzati dati provenienti da diversi archivi amministrativi:

- **Iscritti ai centri per l'impiego:** informazioni sullo stato di disoccupazione e sulla eventuale iscrizione ai corsi di formazione. Anni 2010 e 2011.
- **Comunicazioni Obbligatorie:** dati relativi alle storie lavorative dal 2007 al 2014.
- **Dichiarazioni dei redditi** dell'Agenzia delle Entrate: modelli 730, Unico, 770 dal 2007 al 2013.
- **Ammortizzatori sociali:** dati relativi ai sussidi di disoccupazione erogati dal 2009.
- Dati dell'Agenzia del Lavoro e del Servizio Europa sui **costi** sostenuti per la realizzazione di ogni corso.

Infine per rilevare il titolo di studio si sono utilizzate due banche dati: colloqui Agenzia del Lavoro e Anagrafe Studenti del Servizio Istruzione.

QUALCHE ESEMPIO DI POSSIBILI INTEGRAZIONI DI DATI PER il disegno dei servizi per l'infanzia

Informazioni utili

Caratteristiche dell'utenza potenziale dei servizi socio-educativi per la prima infanzia (3-36mesi)

- Struttura familiare
- Condizioni di vita: livelli di reddito, condizione di deprivazione
- Caratteristiche del luogo di residenza
- Partecipazione al mercato del lavoro dei genitori
- Proprietà casa
- Reti familiari e sociali
- Tempi e soggetti dedicati alla cura dei bambini (se presenti)
- Utilizzo dei servizi per la prima infanzia o di metodi informali di cura
- Motivi del mancato utilizzo
- Livelli di soddisfazione per soluzioni formali di cura dei figli
- Conciliazione tempo lavoro e tempo famiglia e cambiamenti intercorsi pre e post figli

Caratteristiche di chi già usufruisce dei servizi socio-educativi per la prima infanzia (3-36mesi)

- Struttura familiare
- Condizioni di vita: livelli di reddito, condizione di deprivazione
- Tempi e soggetti dedicati alla cura dei bambini
- Livelli di soddisfazione per soluzioni informali di cura dei figli
- Conciliazione tempo lavoro e tempo famiglia e cambiamenti intercorsi pre e post figli

Caratteristiche dell'offerta

- Distribuzione sul territorio e sua evoluzione

Basi di dati

- ✓ Indagini survey sulla popolazione (es. panel, multiscopo Istat, It-Silc)
- ✓ Dati amministrativi relativamente a lavoro (es. INPS, COB), condizioni di vita (es. Agenzia Entrate, APAPI, Archivio Clesius)
- ✓ Indagini survey ah hoc su famiglie con figli in età 3-36m
- ✓ Dati amministrativi sui frequentanti servizi socio-educativi
- ✓ Dati amministrativi su agevolazioni per tariffe
- ✓ Dati amministrativi su presenza asili

QUALCHE ESEMPIO DI POSSIBILI INTEGRAZIONI DI DATI PER il disegno dei servizi di assistenza domiciliare:

Informazioni utili

Caratteristiche dell'utenza potenziale

- Struttura familiare
- Condizioni di vita: livelli di reddito, condizione di deprivazione
- Caratteristiche del luogo di residenza
- Proprietà casa
- Reti familiari e sociali
- Tempi e soggetti dedicati alla cura
- Disagio abitativo
- Salute

Caratteristiche di chi già usufruisce dei servizi

- Caratteristiche degli assistiti
- Natura e consistenza delle prestazioni
- Condizioni di vita: livelli di reddito, condizione di deprivazione
- Livelli di soddisfazione

Caratteristiche dell'offerta

- Caratteristiche e distribuzione sul territorio

Basi di dati

- ✓ Indagini survey sulla popolazione (es. panel, multiscopo Istat, It-Silc)
- ✓ Dati amministrativi relativamente a condizioni di vita (es. Agenzia Entrate, APAPI)
- ✓ Indagini survey ah hoc
- ✓ Dati amministrativi sugli assistiti
- ✓ Dati amministrativi su agevolazioni per tariffe
- ✓ Dati amministrativi

QUALCHE ESEMPIO DI POSSIBILI INTEGRAZIONI DI DATI PER il disegno dei servizi di edilizia abitativa

Informazioni utili

Contesto demografico,
socio-economico

- Struttura delle famiglie
- Disagio abitativo
- Condizioni di vita: livelli di reddito, condizione di deprivazione
- Spese mensili sostenute per la casa
- Caratteristiche del luogo di residenza
- Titolo di godimento dell'abitazione
- Reti familiari e sociali
- Livelli di soddisfazione
- Condizione occupazionale dei componenti adulti
- Caratteristiche eventuale mutuo

Domanda abitativa e
dinamiche del mercato
residenziale

- Evoluzione prezzi e canoni
- Alloggi a canone calmierato
- Dati sul sostegno al canone: graduatorie ed esclusi

Basi di dati

- ✓ Indagini survey sulla popolazione (es. panel, multiscopo Istat, It-Silc)
- ✓ Dati amministrativi relativamente a condizioni di vita (es. Agenzia Entrate, APAPI, Archivio Clesius)
- ✓ Dati bancari su mutui
- ✓ Eventuali indagini ad hoc su partecipanti a progetti sperimentali di social housing
- ✓ Eventuali indagini ad hoc su senza tetto
- ✓ Dati amministrativi sui beneficiari di benefici pubblici
- ✓ Dati amministrativi su agevolazioni per tariffe agevolate